

# Exhibit B



UNITED STATES OF AMERICA  
FEDERAL TRADE COMMISSION  
WASHINGTON, D.C. 20580

Shira A. Steinberg  
Bureau of Competition  
(202) 326-3396  
ssteinberg1@ftc.gov

August 13, 2024

HIGHLY CONFIDENTIAL-ATTORNEYS' EYES ONLY

**BY ELECTRONIC MAIL**

Robert Keeling  
Sidley Austin LLP  
1501 K Street NW  
Washington, DC 20005  
rkeeling@sidley.com

Re: *Federal Trade Commission, et al. v. Amazon.com, Inc.*, No. 2:23-cv-01495-JHC

Dear Robert,

I write in response to Amazon's July 30, 2024 letter regarding TAR 2.0 ("TAR") and its accompanying TAR counterproposal.

**Search Terms and TAR.** As an initial matter, we disagree with your assertion that Plaintiffs' reservations about using TAR in combination with search terms are "unfounded." See Letter from R. Keeling to S. Steinberg at 1 (July 30, 2024) ("July 30 Letter"). While search terms and TAR can be used together, that must be done in a measured manner, accounting for a party's use of both means. Amazon's objections to Plaintiffs' search terms are at odds with the need to use broad search terms here to account for the use of TAR, as Plaintiffs have repeatedly explained to Amazon. See, e.g., Letter from S. Steinberg to R. Keeling (June 10, 2024); Letter from S. Steinberg to K. Trefz and R. Keeling (July 16, 2024); Letter from A. Butler to K. Trefz (July 22, 2024); Letter from S. Steinberg to R. Keeling (July 29, 2024). As you have already acknowledged, "[s]earch terms should be used to identify documents likely to be relevant to the claims and defenses in this litigation," see R. Keeling Letter to S. Steinberg (June 27, 2024), with those documents to be reviewed using TAR. Plaintiffs have proposed appropriate search terms aimed at collecting substantially all documents that may be responsive to Plaintiffs' discovery requests. We understand Amazon will use TAR to facilitate its review of those documents, and we believe Amazon's use of TAR will address any burden concerns.

August 13, 2024

Page 2

Amazon claims that “many” of Plaintiffs proposed search terms “are disconnected from the RFPs and the allegations in Plaintiffs’ Complaint.” *See* July 30 Letter at 1. However, you have identified only two such search strings, even though you have had Plaintiffs’ proposed search terms for nearly three weeks, and despite our express requests that Amazon identify the specific search strings to which it objects.<sup>1</sup> Letter from A. Butler to K. Trefz at 3 (July 22, 2024); Letter from S. Steinberg to K. Trefz at 4,5 (July 29, 2024) Letter from A. Butler to K. Trefz at 2 (Aug. 2, 2024). Plaintiffs are unable to assess Amazon’s complaints about the “many” other search strings that Amazon contends are objectionable given that Amazon has not identified them.

Moreover, despite our repeated express requests that Amazon produce hit reports, including *unique* hits, for any search strings it objects to, *see* Letter from A. Butler to K. Trefz at 3 (July 22, 2024); Letter from S. Steinberg to K. Trefz at 4, 5 (July 29, 2024); Letter from A. Butler to K. Trefz at 2 (Aug. 2, 2024), Amazon has only provided an initial hit count for *all* of Plaintiffs’ initial proposed search terms across the 54 custodians Amazon proposed and what we understand to be non-*unique* hit counts for one search string. *See* Letter from K. Trefz to S. Steinberg et al. at 2 (Aug. 8, 2024). Although Amazon has shown that it has the ability to run hit reports against its proposed list of 54 custodians, you have not substantiated any of your broader objections to Plaintiffs’ initial proposed search terms by providing hit reports, including hit total and *unique* hits for search strings, or any other metric to measure Amazon’s purported burden. Generalized complaints that Plaintiffs’ search terms are overbroad while giving only a few examples are not sufficient and do not comply with the ESI Order in this case. *See* Dkt.# 256 §C.2.c. This is disappointing given Amazon’s repeated commitments to provide hit reports “in response to reasonable requests” in the course of the parties’ negotiations of the ESI protocol. *See, e.g.*, Letter from R. Keeling to E. Bolles at 2 (Jan. 19, 2024); Dkt# 191 at 10 (“Amazon is willing to have the parties substantiate the associated burdens by exchanging hit reports showing counts of documents returned by the terms.

Plaintiffs have agreed to Amazon’s request that—to a certain extent—custodians and search terms be grouped or bucketed together by topic. Our proposal for search terms indicates which search terms relate to the topics of i) market definition and monopoly power, ii) anti-discounting, iii) fulfillment, and iv) general topics<sup>2</sup>. Our proposal for custodians similarly indicates the topics related to each custodian. Amazon has expressed concern with Plaintiffs’ proposal that general topic search terms—such as terms related to competition—should be applied to all custodians. This is an antitrust case, which means that the identity of Amazon’s competitors and how Amazon evaluates and reacts to competitive threats are directly relevant. While Plaintiffs have proposed unique subject-matter strings in certain buckets, generalized strings about competition should be run across all custodians in this competition matter. Amazon has not identified any proposed custodian who bears no responsibility for helping Amazon maintain its competitive position.

---

<sup>1</sup> The only two search strings Amazon has identified are (“Dammit OR Damn OR Darn”), *see* Letter from R. Keeling to A. Butler at 1 (July 30, 2024), and (“Sh\*t OR Shoot” and “F\*ck OR fck OR Fu\*k OR Motherf\* OR WTF\*”), *see* Letter from K. Trefz to S. Steinberg et al. at 2 (Aug. 8, 2024).

<sup>2</sup> In its August 8 letter, Amazon requested Plaintiffs include a fifth bucket for advertising- and search-related custodians and search terms. Plaintiffs will consider this request and respond separately.

August 13, 2024

Page 3

Amazon has additionally expressed concern with Plaintiffs' proposal that all search terms be applied to Amazon's current and former CEOs and current and former S-Team members. Amazon has provided no evidence that applying all search terms to these high-level executives and company leaders will result in a high number of false positives.<sup>3</sup> To the extent that this limited set of custodians share copies of the same document, this will be resolved by de-duplication and thus there will be no increased burden. Moreover, the S-Team is a cross-functional leadership team, reporting to the CEO, with oversight responsibilities across Amazon's business lines. At Amazon, leaders "act on behalf of the entire company, beyond just their own team. . . [who] operate at all levels, [and] stay connected to the details."<sup>4</sup> The appropriate bucket of search terms for CEOs and S-Team is all search terms.

**TAR 2.0.** Amazon largely ignored Plaintiffs' July 16 TAR proposal and rejected most of Plaintiffs' proposed edits with no explanation. It is disappointing that after three meet and confers and exchanging multiple letters to date, *see e.g.*, Letter from S. Steinberg to R. Keeling (June 10, 2024); Letter from R. Keeling to S. Steinberg (June 27, 2024); Letter from S. Steinberg to K. Trefz and R. Keeling (July 16, 2024), Amazon appears unwilling to consider any substantive elements of Plaintiffs' proposal in its TAR protocol.<sup>5</sup> Amazon appears to have disregarded Plaintiffs' proposal, even in instances where Amazon previously agreed to these changes. We believe the parties are in agreement on the following provisions and have accordingly reinserted them in the proposed TAR protocol attached to this letter. If Amazon disagrees with any of the following sections, please confirm and explain your basis for removing each from the TAR protocol:

- To ensure the parties have a full understanding of how TAR will be used, Plaintiffs included a provision requiring Amazon "to disclose custodians or repositories Amazon plans on reviewing in a linear fashion outside of TAR prior to finalizing the TAR protocol." July 30 Letter at 7. You agreed to this in your July 30 Letter.
- During our call on July 3, Plaintiffs proposed email threading be adopted to limit the number of documents, and in turn Amazon's burden, to review documents. *See* Letter from S. Steinberg to R. Keeling, Appendix A at § 3.4 (July 16, 2024). You acknowledged threading was possible during the parties' call on July 9 and agreed to adopt threading during the same call. Given that email threading will

---

<sup>3</sup> Amazon's July 30 Letter claims the "parties agreed to apply search terms for RFPs 6 and 44 to all of Amazon's proposed 54 custodians." This is not correct. Plaintiffs view RFP 44 as a request aimed at market definition and monopoly power custodians, though Plaintiffs will not object if Amazon treats RFP 44 as a general RFP applicable to all custodians.

<sup>4</sup> *Leadership Principles*, Amazon <https://www.amazon.jobs/content/en/our-workplace/leadership-principles>.

<sup>5</sup> The revisions to the TAR protocol attached to Amazon's July 30 letter are titled "Redline – Amazon's Second TAR Proposal vs. Amazon's First TAR Proposal." However, Amazon appears to have edited the "redline" proposed TAR protocol provided along with its July 30 Letter to include some of Plaintiffs' proposed edits not in track changes. *See* Amazon's July 30 Redline at 6.3. Plaintiffs assume this is an oversight. Please provide a redline showing *all* differences between Amazon's first TAR proposal and its second TAR proposal by **August 16**. Or, as is done in the normal exchange of proposals, Amazon should commit to providing red lines to the most recent version of the TAR protocol proposed instead of editing its own proposal partially in redline and partially with changes incorporated not in redline.

August 13, 2024

Page 4

greatly reduce Amazon's cost and burden to review, we do not understand why Amazon would remove this section from its proposed TAR protocol.

- Plaintiffs proposed that Amazon remove documents identified as “spam and junk” before such documents are loaded into Brainspace. Plaintiffs initially raised this question with Amazon during our July 3 call. During our July 9 call, Amazon agreed to remove spam and junk from the document population before ingesting into Brainspace. Plaintiffs do not understand why Amazon refused to incorporate this proposed measure to the TAR protocol, particularly because removing spam and junk from the document population is a standard part of the document review process, aimed at lowering Amazon’s overall burden and reducing the total number of documents Amazon would have to review.

You deleted nearly all of Plaintiffs’ proposed changes with no explanation in your counterproposal or accompanying letter. We have added these provisions back in for Amazon’s review and in the hopes of furthering the parties’ discussion of these sections. For each, Plaintiffs would like to understand the basis for your removal in the hopes of reaching agreement on a TAR protocol that is acceptable to all parties. For each provision Amazon previously deleted, please confirm whether Amazon will agree to Plaintiffs’ proposal and if not, explain why by **August 20:**

- Plaintiffs proposed limiting the extracted text file size of documents loaded into Brainspace to under 1 megabyte. Plaintiffs’ suggestion is in line with Brainspace’s own recommendation that “the general best practice is to filter out any document that exceeds an extracted text size of 1MB.” Brainspace, Pre-Filtering Data, <https://brainspace-help.revealdata.com/en/Pre-filtering-Data.html>. Plaintiffs’ suggestion is reasonable and should be noncontroversial, particularly since Brainspace recommends “the lower [extracted text size] the better.” *Id.*
- Plaintiffs proposed including a provision, *see* Appendix A at §3.6, wherein if Amazon uses bulk coding in its review, it will provide Plaintiffs with basic information regarding how such documents were identified and determined to be suitable for bulk coding. This is a reasonable step to ensure Amazon is not inadvertently removing documents from the population via bulk coding that may be necessary to train the model. To be clear, Plaintiffs do not believe bulk coding is appropriate once documents have been ingested in Brainspace. We are willing to forego this provision if Amazon confirms it does not plan to use bulk coding as part of its review.
- You agreed to provide recall, richness, precision, and elusion rates as part of Plaintiffs’ validation, but wholesale deleted Plaintiffs’ proposed method of calculating these rates. To be clear, Plaintiffs’ proposed definitions are based off of the industry standard for calculating these metrics and the language was derived from Relativity’s definitions for calculating each rate. *See* Relativity, Project Validation Statistics, [https://help.relativity.com/RelativityOne/Content/Relativity/Active\\_Learning/Project\\_Validation\\_Statistics.htm](https://help.relativity.com/RelativityOne/Content/Relativity/Active_Learning/Project_Validation_Statistics.htm). If Amazon is unwilling to use Plaintiffs’ definition

August 13, 2024

Page 5

for these rates, please confirm how Amazon plans to calculate recall, richness, precision, and elusion rates.

- You removed Plaintiffs' edits regarding Plaintiffs' method of validating Amazon's TAR review. *See Appendix A at §7.2.* Amazon proposed providing Plaintiff reviewers with access to a separate database for 7 days to review the sample; Plaintiffs' edits, instead, requested Amazon produce the documents to Plaintiffs and permit Plaintiffs 14 days to review the sample. The parties agree that after Plaintiffs review, the parties will meet and confer to discuss "whether there is good reason to believe that the recall level has not been achieved." Amazon suggests Plaintiffs need to use a separate database for its review because of "security reasons." We are unsure why this would be a concern given the parties' expansive Protective Order, Dkt.# 160, governing the parties' use of much more confidential and sensitive information than what would be expected in a sample set of non-responsive documents for Plaintiffs' validation.

In response to many of Plaintiffs' questions, you pointed to your prior experience responding to HSR Second Requests from the FTC or Department of Justice. We do not agree that Second Requests are a useful guidepost for the use of TAR in litigation. Among other issues, Second Requests are part of the agencies' merger investigations, and as we have previously explained, investigations serve a different purpose than discovery in litigation. *See Joint Status Report at 28-29, Dkt. #246.*

During the parties' negotiations, Amazon has asserted that its TAR model, Brainspace, cannot ingest more than 9 million documents or it will "become unstable." July 30 Letter at 2; *see also* Letter from K. Trefz to S. Steinberg at 2 (Aug. 8, 2024). Plaintiffs have sought to understand Amazon's basis for this belief, particularly since Plaintiffs have never encountered such a limitation despite our familiarity with Brainspace. *See* July 30 Letter at 5. Plaintiffs have previously spoken to a sales representative from Brainspace and asked for information on whether there is a limit on the number of documents Brainspace can ingest. Contrary to Amazon's assertions, Brainspace has represented that it frequently has clients feeding 70 to 80 million documents into the program with no issue and they have seen matters where *billions* of documents are reviewed using Brainspace. From speaking with Brainspace, Plaintiffs understand that Brainspace itself does not have performance limitations that would be relevant here. Rather, any limit on the program's performance would arise from the server space or capacity allotted to Brainspace by the client or its vendor (here, Amazon and Consilio).

Based on your explanations,<sup>6</sup> we understand Amazon is concerned that if Brainspace processes over 9 million documents, it "may not be able to perform efficiently" and will take time to process documents. In support of the second point, you report that as the model is "updated based on newly available coding decisions . . . it takes longer and longer to index the

---

<sup>6</sup> You cite to Relativity's recommended "performance baseline" to support your opinion that "Brainspace can become unstable when applied to more than about 9 million documents." However, on the same page Amazon cites to, Relativity offers options for parties with a larger data set which explains how to utilize "Active Learning across tens of millions of documents." *See* Relativity, Scaling Active Learning, [https://help.relativity.com/RelativityOne/Content/Solutions/Scaling\\_Active\\_Learning.htm](https://help.relativity.com/RelativityOne/Content/Solutions/Scaling_Active_Learning.htm)

August 13, 2024

Page 6

documents.” However, indexing will progressively take more time as more documents are loaded onto the platform over the course of Amazon’s review. Indexing will not pose a problem when there are fewer documents in the model. Accordingly, the time needed to index documents should only be a concern for a small portion of Amazon’s review: the point at which the ultimate volume of documents has been loaded into the model. It is also not clear to us how much time, if any, will be lost as a result of Amazon needing to reindex its model in the course of its review. Indexing will only take “several days” at the tail end of Amazon’s review when it finally loads the remaining documents (which may or may not surpass nine million documents) into the model. Amazon also claims that it may take “several days” for indexing to be complete because the model may “error out” partway through and the process must “be restarted.” July 30 Letter at 3. However, Brainspace’s sales representative confirmed that if the model does “error out,” the model’s progress reindexing will be saved up to the point the model crashed. Additionally, in Plaintiffs’ experience, reviewers are able to continue working while the model is indexing, so there is no delay in reviewing documents while indexing occurs. As a result, our understanding is that Amazon can continue reviewing documents while the model reindexes, even if that reindexing proceeds in parts.

To the extent Amazon continues to be concerned about the efficiencies and time required to use Brainspace, we believe Amazon’s concerns can be readily addressed in a number of ways, as we offered in our July 16 letter. *See Letter from S. Steinberg to R. Keeling at 3 (July 16, 2024).* First, once a large number of documents have been loaded into Brainspace, if Amazon should experience delays due to reindexing, Amazon could choose to index its model less frequently (*i.e.* every few days or once a week) to lessen any delays. Again, Plaintiffs do not view this as a delay because reviewers can continue coding while the index is being rebuilt. Second, Plaintiffs previously informed Amazon that “TAR models generated by Brainspace are portable, such that they can be applied to more than one dataset in parallel.” You claimed that this feature “has no bearing on Brainspace’s ability to scale beyond 9 million documents,” but that is not correct. If Amazon cannot manage to use Brainspace with more than 9 million documents, it can use Brainspace to review an initial 9 million documents, export the trained TAR model into a new database, and use that database to review additional documents. Brainspace has information available about this process on its website that Plaintiffs previously cited to, *see* Brainspace, Portable Models, <https://brainspace-help.revealdata.com/en/Portable-Models.html>. Here, Brainspace confirms the training model can be loaded into a new database and “used just as if it was trained on that dataset. It can also be updated with training data from the dataset to further improve it.” *Id.* When Plaintiffs spoke with Brainspace’s sales representative, he also confirmed this was a possibility. Plaintiffs have used similar methods of importing models in the past and are willing to speak with Amazon to explain how this method can be employed here.

Given that Amazon has “one of the country’s preeminent authorities on eDiscovery” managing discovery<sup>7</sup> and is employing Consilio, “one of the largest, most recognized, and most respected eDiscovery vendors worldwide,” *see* July 30 Letter at 3, we believe Amazon is more than capable of resolving these issues and concerns. Regardless, the parties currently do not know the ultimate number of documents that will be identified as potentially containing

---

<sup>7</sup>*See* Robert D. Keeling, Sidley Austin, <https://www.sidley.com/en/people/k/keeling-robert-d>,

August 13, 2024

Page 7

responsive material based on search terms and it is unclear if this will truly pose a concern for the parties at all. To be clear, Plaintiffs do not believe that the purported 9 million document limit is a basis to cap Plaintiffs' discovery.

Plaintiffs have also suggested Amazon exclude standalone spreadsheets from its TAR review, as this document format is often inaccurately coded using TAR. As you know, spreadsheets will be analyzed by TAR through its extracted text. It remains unclear why the TAR model would be better able to discern the relevancy of such spreadsheets that would be "likely meaningless for human reviewers." You suggest that if certain spreadsheets "contain too little or too much content" they will be removed from Brainspace. Plaintiffs, again, propose adopting a limit on the extracted text file size to 1 megabyte as a way of easily delineating which files contain too much text and providing Plaintiffs some assurance that spreadsheets will be appropriately reviewed. Please confirm your acceptance or rejection of this proposal by **August 16**.

With respect to recall rate, Plaintiffs have now offered two approaches for determining an acceptable recall rate. Plaintiffs first suggested Amazon adopt a method of calculating recall that accounts for the use of search terms and TAR by considering the number of responsive documents left in the overall document population before it was culled by search terms and TAR. *See Letter from S. Steinberg to R. Keeling (June 10, 2016) at 3; see also Letter from S. Steinberg to K. Trefz and R. Keeling at 5 (July 16, 2024).* Amazon never responded to this proposal, either during the parties July 3 call where Plaintiffs first proposed it nor in Amazon's July 30 letter. Plaintiffs next suggested Amazon proceed with a 75% recall rate unless either the precision or richness rate was over 50%, in which case Amazon would continue reviewing documents until an 85% recall level was reached. In contrast, your proposal of a 75% recall rate has not changed since your initial proposal on May 28 and you have offered no other counterproposals or attempts to compromise.<sup>8</sup> Our proposal is aimed at tailoring the TAR protocol to the facts of the case, here, rather than an abstract recall rate with no basis, as Amazon proposed. As you are aware, "the facts and circumstances of each case are different, and no rigid standards are appropriate." Bolch Judicial Institute, *Technology Assisted Review (TAR) Guidelines* at 23 available at <https://edrm.net/wp-content/uploads/2019/02/TAR-Guidelines-Final.pdf> (R. Keeling listed as contributor). Adopting Plaintiffs' recall method would apply a flexible recall rate that increases if richness or precisions, two metrics aimed at estimating the proportion of documents in a data set that are relevant, are above 50%. Our proposal relies on the basic tenant that if it is more likely than not (*i.e.*, a greater than 50% chance) that Amazon can identify responsive documents in the population, it should continue doing so until approximately 85% of the relevant documents are found in the population. This proposal also builds on this paper's guidance "that achieving a recall rate of 75% to 85% has been a good balance in many cases." *Id.*

---

<sup>8</sup> Instead, Amazon has also threatened, twice, to adopt a "traditional, search-terms-only approach without the overhead of a TAR workflow" if Plaintiffs insist on a higher recall rate. *See Letter from R. Keeling to S. Steinberg at 7 (July 30, 2024); Letter from R. Keeling to S. Steinberg at 4 n.2 (June 27, 2024).* Plaintiffs have no objection in principle to Amazon using only search terms. However, any unwarranted delay as the result of Amazon changing its proposed review method will be promptly raised with the court.

August 13, 2024

Page 8

Amazon objects to Plaintiffs' proposal due to "the cost and burden of achieving higher recall levels" and you suggest if an 85% recall rate were adopted, Amazon is "better off pursuing a traditional, search-terms-only approach without the overhead of a TAR workflow." July 30 Letter at 7. This objection is unfounded. As we have previously noted, Amazon has provided no evidence of its purported undue burden. If Amazon believes it is able to estimate the burden of achieving a recall rate of 85% rather than 75%, please provide any metrics or evidence to support your objection. Without information to the contrary, Plaintiffs maintain that, even with a higher recall rate, Amazon's burden is greatly reduced by using TAR because it does not have to review every document for relevance and the parties can use broader search terms to streamline those negotiations. The parties can also save time on negotiating search terms because a TAR review will separately winnow out the documents least likely to contain responsive materials. If Amazon continues to disagree with Plaintiffs' proposal, Amazon can make a counterproposal of its own. To ensure the parties are moving forward with finalizing the TAR protocol in a timely manner, please provide a response by **August 20**.

Plaintiffs have additional concerns about Amazon using TAR when the applicable date range differs for certain categories of documents. Amazon currently continues to state that a January 1, 2018 default start date is appropriate,<sup>9</sup> but agreed to produce documents beginning January 1, 2015 for SFP-related documents, and to produce documents beginning January 1, 2014 for Nessie-related documents. With respect to SFP, Amazon explained it plans to collect SFP-related materials for the date range by running SFP-related search terms for certain custodians, and that "reviewers will be trained to mark responsive SFP-related materials that fall within the date range." Email from K. DeWilde to C. Herd and M. Baker (July 26, 2024). Plaintiffs are concerned that Amazon marking non-SFP related materials as non-responsive based solely on date will improperly train the TAR model to categorize those documents as substantively not relevant. We are also concerned this is not the most efficient path forward as limiting discovery in this manner inevitably increases Amazon's own burden by adding non-responsive documents to Brainspace and giving additional documents for its reviewers to work through. Please explain how Amazon plans to treat non-SFP (or Nessie) related materials that fall outside the agreed upon default date range so that the TAR model is still properly trained. Please provide your explanation by **August 20**.

**Document Collection.** We understand Amazon has collected, at a bare minimum, custodial data from "backend sources (e.g., email, Chime, etc.)" for all of Amazon's proposed custodians. However, Plaintiffs' question was broader and requested Amazon identify which custodians' files Amazon has collected to date, which collections are in progress, and to let us know by what date Amazon plans to have completed custodial collections for all custodians in its June 27 proposal. Additionally, if you have identified any custodians that do not have custodial documents, please provide an update for Plaintiffs as soon as possible as this information is necessary for the parties' custodian negotiations. Plaintiffs respectfully request, again, that you provide this information by **August 20**.

**Non-Custodial Data Sources.** Amazon's positions regarding non-custodial data sources raise a number of concerns.

---

<sup>9</sup> Plaintiffs continue to maintain that a default January 1, 2014 start date is appropriate.

August 13, 2024

Page 9

First, the ESI Order required Amazon to disclose “[a] list of non-custodial data sources (e.g., shared drives, servers), if any, likely to contain discoverable ESI.” Dkt.# 256 at § B(2). In our July 16 letter we noted that Amazon’s June 20 ESI disclosure of non-custodial data sources, which was wholly based on its February 22 letter, was seriously deficient. Your July 30 letter explains that Amazon’s response did not distinguish between “non-custodial data sources likely to contain responsive information” and “other non-custodial data sources that we do not believe to be relevant.” July 30 Letter at 8. By your own admission, Amazon’s June 20 ESI disclosures are deficient, and frankly of little use for the purpose of meaningfully negotiating the scope of Amazon’s search for responsive documents. We reiterate our request that you provide a “list of non-custodial data sources (e.g., shared drives, servers), if any, likely to contain discoverable ESI” by **August 20**, or explain Amazon’s refusal to do so.

Second, you categorize several data [REDACTED]

[REDACTED] as “custodial” rather than “non-custodial.” July 30 Letter at 8. In your August 7 letter, you also refer to Salesforce as containing “custodial Salesforce documents.” As an initial matter, the ESI Order—in provisions Amazon agreed to—specifically refers to “shared drives” as a non-custodial source. Dkt.# 256 at § B(2); *see also* Dkt.# 191-1 at § B(2) (Amazon’s proposed ESI Order). As a result, we believe it is clear that all of these data sources, which are analogous to shared drives, should be treated as non-custodial document repositories. Further, so that we can ascertain the extent of our disagreement, please explain how Amazon proposes to search for and collect documents from these sources. For example only, it would be insufficient, as Amazon proposed for [REDACTED] in your August 7 letter, for Amazon to collect only files in a shared repository where a custodian was an author or input comments; Amazon should fully collect or run search terms against any repository where a custodian had access to documents in the course of their work.

Third, it is not clear to us whether Amazon is treating any other data sources identified in its June 20 disclosure as custodial, rather than non-custodial, repositories. Aside from [REDACTED], please identify any data sources in Amazon’s June 20 that Amazon is classifying as “custodial” sources by **August 20**.

Fourth, we are concerned that Amazon has not made a diligent effort to identify non-custodial sources likely to contain information responsive to Plaintiffs’ discovery requests. Amazon’s June 20 ESI disclosure, as noted above and in our July 16 letter, is seriously deficient, and does not provide the information that would allow the parties to meaningfully discuss Amazon’s collection, review, and production of documents from non-custodial sources. [REDACTED]  
[REDACTED]  
[REDACTED]  
[REDACTED]

Amazon must conduct a reasonable search for and of non-custodial sources, and as we noted in our July 16 letter, we have repeatedly confirmed our position that Amazon must produce documents from non-custodial as well as custodial sources. *See, e.g.*, Letter from M. Baker to K. Trefz at 2 n. 2 (June 4, 2024). If Amazon does not agree that it is required to conduct a

August 13, 2024

Page 10

reasonable search for and of non-custodial sources for documents responsive to Plaintiffs' discovery requests, please confirm that position and explain the basis for it by **August 20**. If Amazon agrees that it is required to conduct a reasonable search for and of non-custodial sources, please explain by **August 20**: (a) what Amazon has done so far to search for non-custodial sources; (b) what Amazon is continuing to do to search for non-custodial sources; (c) what non-custodial sources Amazon is searching or commits to search for responsive documents; and (d) how Amazon is conducting those searches, in light of your statement that Amazon does not plan to search any non-custodial document repositories using search terms.

**Timing.** Plaintiffs proposed a set of milestones for discovery that were reasonable and aimed at ensuring Amazon is making progress in line with the Court's Scheduling Order, Dkt. #159. When Plaintiffs initially raised this schedule during our July 3 call, we emphasized that "the parties should work together to set milestones" and requested that if Amazon did not find the schedule workable, it provide a counterproposal in a timely manner. *See* July 16 Letter at 9. Despite having our proposal in hand for over a month, Amazon has not provided a counterproposal or any sense of its estimated timing to substantially complete document production.

The timeline set out by Plaintiffs was wholly feasible but for Amazon's continued delays. By way of example, the parties exchanged search terms on July 22 and Plaintiffs provided their initial custodian proposal on July 25. After nearly three weeks, Amazon made a broad claim of undue burden, but has yet to provide hit counts or substantive objections. *See* Letter from K. Trefz to S. Steinberg et al. at 2 (Aug. 8, 2024). Similarly, it took Amazon over two weeks to respond to Plaintiffs' July 16 letter containing easily answerable questions about Amazon's use of TAR. And Amazon has already raised concerns about the time needed to complete its document review both with the court, *see* Dkt. #246 at 24, and again in its July 30 Letter. We continue to hope that the parties can work together to create a reasonable timeline to ensure that Amazon makes steady progress and substantially completes its document production by the end of 2024. Such a schedule would enable the parties to move forward with depositions and conduct any necessary follow-up discovery in 2025. Accordingly, please provide a counterproposal to Plaintiffs' proposed milestones by **August 20**.

\* \* \*

We are available to meet and confer regarding any of the topics discussed above.

Sincerely,

/s/ Shira A. Steinberg  
Shira A. Steinberg  
Attorney  
Bureau of Competition  
Federal Trade Commission